Graph Theory

Linear algebra background

Instructor: *Nicolò Cesa-Bianchi* version of June 3, 2020

The material in this handout is taken from: Luca Trevisan, Lecture Notes on Graph Partitioning, Expanders and Spectral Methods, 2016.

Given a real $n \times n$ matrix M, if $M\boldsymbol{u} = \lambda \boldsymbol{u}$ for some $\lambda \in \mathbb{R}$ and $\boldsymbol{u} \in \mathbb{R}^d \setminus \{\boldsymbol{0}\}$, then $\boldsymbol{u}$ is an eigenvector of M with eigenvalue λ (we also say that $\boldsymbol{u}$ is an eigenvector of λ). Note that eigenvectors can be rescaled without changing the equation $M\boldsymbol{u} = \lambda \boldsymbol{u}$, hence we conventionally assume they have unit length.

Note that λ is an eigenvalue for M if and only if there exists $\boldsymbol{x} \neq \boldsymbol{0}$ such that $(M - \lambda I)\boldsymbol{x} = \boldsymbol{0}$, where I is the $n \times n$ identity matrix. The equation $(M - \lambda I)\boldsymbol{x} = \boldsymbol{0}$ holds for $\boldsymbol{x} \neq \boldsymbol{0}$ if and only if $M - \lambda I$ is singular, which is equivalent to $\det(M - \lambda I) = 0$. Since $\det(M - \lambda I)$ is a n-th degree univariate polynomial in λ, it has exactly n solutions by the fundamental theorem of algebra. This shows that every square matrix has n eigenvalues (not all necessarily distinct). Some of these eigenvalues, however, may correspond to solutions of $\det(M - \lambda I) = 0$ in the complex plane. The next result guarantees that at least one eigenvalue is real when M is symmetric.

Fact 1 (proof omitted) *If M is symmetric, then there exists $\lambda \in \mathbb{R}$ and $\boldsymbol{u} \in \mathbb{R}^n \setminus \{\boldsymbol{0}\}$ such that $M\boldsymbol{u} = \lambda \boldsymbol{u}$.*

Fact 2 *If M is symmetric then any two eigenvectors corresponding to distinct eigenvalues are orthogonal.*

Proof. Let $\boldsymbol{x}$ be an eigenvector of λ and $\boldsymbol{y}$ an eigenvector of λ'. Since M is symmetric, $(M\boldsymbol{x})^\top \boldsymbol{y} = \boldsymbol{x}^\top M \boldsymbol{y}$. On the other hand, $(M\boldsymbol{x})^\top \boldsymbol{y} = \lambda \boldsymbol{x}^\top \boldsymbol{y}$ and $\boldsymbol{x}^\top M \boldsymbol{y} = \lambda' \boldsymbol{x}^\top \boldsymbol{y}$. Since $\lambda \neq \lambda'$, it must by $\boldsymbol{x}^\top \boldsymbol{y} = 0$, which means that $\boldsymbol{x}$ and $\boldsymbol{y}$ are orthogonal. □

Theorem 3 (Spectral Theorem) *Let $M \in \mathbb{R}^{n \times n}$ be a real symmetric matrix. Then there exists n (not necessarily distinct) real numbers $\lambda_1, \dots, \lambda_n$ and n orthonormal real vectors $\boldsymbol{u}_1, \dots, \boldsymbol{u}_n$ such that $\boldsymbol{u}_i$ is an eigenvector of λ_i.*

Proof. The proof is by induction on n. If $n = 1$, then M is a scalar. Hence, $Mx = Mx$ shows that any nonzero $x \in \mathbb{R}$ is an eigenvector of M with eigenvalue M.

Assume now that the statement holds for $n - 1$. By Fact 1, there exist an eigenvalue $\lambda_1 \in \mathbb{R}$ with eigenvector $\boldsymbol{x}_1 \in \mathbb{R}^n$. Now we claim that $\boldsymbol{y}$ orthogonal to $\boldsymbol{x}_1$ implies $M\boldsymbol{y}$ is orthogonal to $\boldsymbol{x}_1$. Indeed, $\boldsymbol{x}_1^\top M \boldsymbol{y} = (M\boldsymbol{x}_1)^\top \boldsymbol{y} = \lambda \boldsymbol{x}_1 \boldsymbol{y} = 0$.

Let V be the $n - 1$-dimensional subspace of $\mathbb{R}^n$ that contains all the vectors orthogonal to $\boldsymbol{x}_1$. We can define a one-to-one linear mapping $B : \mathbb{R}^{n-1} \to V$ by choosing an orthonormal basis $\boldsymbol{u}_1, \dots, \boldsymbol{u}_{n-1}$ for V and letting $B = \big[\boldsymbol{u}_1, \dots, \boldsymbol{u}_{n-1}\big]$. Also, $BB^\top$ is a projection matrix that projects vectors from $\mathbb{R}^n$ to V. In particular, $BB^\top \boldsymbol{y} = \boldsymbol{y}$ for all $\boldsymbol{y} \in V$. We now apply the inductive hypothesis to the $(n-1)\times(n-1)$ matrix $M' = B^\top MB$ and find real eigenvalues $\lambda_1, \dots, \lambda_{n-1}$ and orthonormal eigenvectors $\boldsymbol{y}_1, \dots, \boldsymbol{y}_{n-1}$. For $2 = 1, \dots, n$ we have $M'\boldsymbol{y}_i = B^\top MB\boldsymbol{y}_i = \lambda_i \boldsymbol{y}_i$. Therefore, $BB^\top MB\boldsymbol{y}_i = \lambda_i B\boldsymbol{y}_i$. Since $\boldsymbol{y}_i \in \mathbb{R}^{n-1}$ and B maps $\mathbb{R}^{n-1}$ to V, $B\boldsymbol{y}_i$ is orthogonal to $\boldsymbol{x}_1$ and, by the above claim, $MB\boldsymbol{y}_i$ is orthogonal to $\boldsymbol{x}_1$. Therefore $BB^\top MB\boldsymbol{y}_i = MB\boldsymbol{y}_i$ and so we have $MB\boldsymbol{y}_i = \lambda_i B\boldsymbol{y}_i$. If we now define $\boldsymbol{x}_i = B\boldsymbol{y}_i$ we have $M\boldsymbol{x}_i = \lambda_i \boldsymbol{x}_i$. To finish up, note that by construction $\boldsymbol{x}_1$ is orthogonal to $\boldsymbol{x}_2, \dots, \boldsymbol{x}_n$. Moreover, for any $2 \le i < j \le n$, $\boldsymbol{x}_i^\top \boldsymbol{x}_j = (B\boldsymbol{y}_i)^\top (B\boldsymbol{y}_j) = \boldsymbol{y}_i^\top B^\top B\boldsymbol{y}_j = \boldsymbol{y}_i^\top \boldsymbol{y}_j = 0$. Hence we have found n eigenvalues with n eigenvectors. □

Corollary 4 *Let $M \in \mathbb{R}^{n\times n}$ be a real symmetric matrix. Then*

$$M = U\Lambda U^\top = \sum_{i=1}^{n} \lambda_i \boldsymbol{u}_i \boldsymbol{u}_i^\top$$

where $U = \big[\boldsymbol{u}_1, \dots, \boldsymbol{u}_n\big]$ and $\Lambda = \operatorname{diag}\big(\lambda_1, \dots, \lambda_n\big)$. Here $\lambda_1 \le \lambda_2 \le \dots \le \lambda_n$ are the real eigenvalues of M and $\boldsymbol{u}_1, \dots, \boldsymbol{u}_n \in \mathbb{R}^n$ are the corresponding eigenvectors.

Proof. Note that $MU = \big[\lambda_1\boldsymbol{u}_1, \dots, \lambda_n\boldsymbol{u}_n\big]$ because $M\boldsymbol{u}_i = \lambda_i\boldsymbol{u}_i$ for each $i = 1, \dots, n$. Hence $MU = U\Lambda$ where $\Lambda = \operatorname{diag}\big(\lambda_1, \dots, \lambda_n\big)$. Since U is orthonormal, $U^{-1} = U^\top$ and $UU^\top = I$. Therefore $M = MUU^\top = U\Lambda U^\top$. □

Theorem 5 (Variational characterization of eigenvalues — proof omitted) *Let $M \in \mathbb{R}^{n\times n}$ be a real symmetric matrix, and $\lambda_1 \le \lambda_2 \le \dots \le \lambda_n$ be its real eigenvalues. For $k < n$ let $\boldsymbol{u}_1, \dots, \boldsymbol{u}_k$ be orthonormal vectors such that $M\boldsymbol{u}_i = \lambda_i\boldsymbol{u}_i$ for $i = 1, \dots, k$. Then*

$$\lambda_{k+1} = \min_{\substack{\boldsymbol{u}\in\mathbb{R}^n\setminus\{\boldsymbol{0}\}\\ \boldsymbol{u}\perp\{\boldsymbol{u}_1,\dots,\boldsymbol{u}_k\}}} \frac{\boldsymbol{u}^\top M\boldsymbol{u}}{\boldsymbol{u}^\top\boldsymbol{u}}$$

and any minimizer $\boldsymbol{u}$ is an eigenvector of λ_{k+1}.

The ratio in the right-hand side is called Rayleigh quotient. Note that, in particular,

$$\lambda_1 = \min_{\boldsymbol{u}\in\mathbb{R}^n\setminus\{\boldsymbol{0}\}} \frac{\boldsymbol{u}^\top M\boldsymbol{u}}{\boldsymbol{u}^\top\boldsymbol{u}} \ .$$

Also, because $-M$ has eigenvalues $-\lambda_n \le -\lambda_{n-1} \le \dots \le -\lambda_1$,

$$-\lambda_n = \min_{\boldsymbol{u}\in\mathbb{R}^n\setminus\{\boldsymbol{0}\}} \frac{\boldsymbol{u}^\top(-M)\boldsymbol{u}}{\boldsymbol{u}^\top\boldsymbol{u}} = -\max_{\boldsymbol{u}\in\mathbb{R}^n\setminus\{\boldsymbol{0}\}} \frac{\boldsymbol{u}^\top M\boldsymbol{u}}{\boldsymbol{u}^\top\boldsymbol{u}}$$

and therefore

$$\lambda_n = \max_{\boldsymbol{u}\in\mathbb{R}^n\setminus\{\boldsymbol{0}\}} \frac{\boldsymbol{u}^\top M\boldsymbol{u}}{\boldsymbol{u}^\top\boldsymbol{u}}$$